

Frequency-aware optical coherence tomography image super-resolution via conditional generative adversarial neural network

XUESHEN LI,¹ ZHENXING DONG,²  HONGSHAN LIU,¹ JENNIFER J. KANG-MIELER,¹ YUYE LING,²  AND YU GAN^{1,*} 

¹*Department of Biomedical Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA*

²*Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, Minhang District, 200240, China*

*ygan5@stevens.edu

Abstract: Optical coherence tomography (OCT) has stimulated a wide range of medical image-based diagnosis and treatment in fields such as cardiology and ophthalmology. Such applications can be further facilitated by deep learning-based super-resolution technology, which improves the capability of resolving morphological structures. However, existing deep learning-based method only focuses on spatial distribution and disregards frequency fidelity in image reconstruction, leading to a frequency bias. To overcome this limitation, we propose a frequency-aware super-resolution framework that integrates three critical frequency-based modules (i.e., frequency transformation, frequency skip connection, and frequency alignment) and frequency-based loss function into a conditional generative adversarial network (cGAN). We conducted a large-scale quantitative study from an existing coronary OCT dataset to demonstrate the superiority of our proposed framework over existing deep learning frameworks. In addition, we confirmed the generalizability of our framework by applying it to fish corneal images and rat retinal images, demonstrating its capability to super-resolve morphological details in eye imaging.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Optical coherence tomography (OCT) is a non-invasive imaging modality that utilizes infrared interferometry to generate depth-resolved reflectivity profile in real-time [1]. Over the last decades, OCT has stimulated a wide range of medical image-based diagnosis and treatment [2–4]. For example, in cardiology, OCT is considered a suitable coronary imaging modality to assess plaques to ensure successful stent deployment [5]. Meanwhile, in ophthalmology, OCT has become one of the prominent diagnostic tools for keratoconus [6], glaucoma [7], age-related macular degeneration [8], retinopathy [9], diabetic retinopathy, and diabetic macular edemas [10] in identifying layers in both anterior and posterior segments of the eye.

In both coronary imaging and eye imaging, high spatial resolutions from OCT, mostly spectral domain OCT (SDOCT), are crucial to facilitate the application to either identify endothelial cells or assess the thickness of corneal layers and retinal layers. However, such high resolution comes at the cost of demanding optical design and data transmission/storage. Improvement of resolution via upgrading light sources and other hardware designs is resource-intensive. The hardware improvement also suffers from jittering and motion artifacts caused by sparse sampling. On the contrary, software-based method could bypass the hardware upgrading issue and achieve high image quality through computational methods.

In the realm of algorithmic super-resolution (SR), various digital signal processing and image processing methods, have been developed to generate high-resolution (HR) OCT images from low-resolution (LR) OCT scanning which is undersampled in spectral or/and spatial domain. Conventionally, deconvolution [11,12], spectrum-shaping [13], and spectral estimation [14] have

been proposed to optimize OCT images. However, the SR performance is recently boosted by the introduction of deep learning (DL), especially the combination of convolution neural network (CNN) and generative adversarial network (GAN).

Convolution neural network has been widely used for OCT image generation [15–22] to enhance the image quality and denoise speckles [1,17,23–25]. However, previous CNN models are not frequency-aware. In particular, CNN models such as multi-scale residual network (MSRN), residual dense network (RDN), residual dense Unet (RDU), and Residual Channel Attention Network (RCAN) have been utilized and compared recently in generating SR OCT images [21]. Moreover, conditional generative adversarial network (cGAN) has been incorporated in OCT SR research [17,26–29], featuring a discriminator design to examine the fidelity of the generated SR image during the training process, thus enhancing the capability of generating HR images.

However, the current DL research in generating SR OCT images focuses solely on the spatial distribution of pixels in B-scans, without consideration of frequency information. The lack of frequency-awareness poses limitation on SR performance in two aspects. Firstly, from 1-D frequency perspective, SDOCT is physically measured in spectrum and reconstructed in spatial domain. Considering frequency information along axial direction would increase the fidelity of reconstruction. Secondly, from 2-D image processing perspective, current DL models exhibit spectral bias, which is a learning bias towards low-frequency components [30,31]. As shown in Fig. 1, DL algorithms induce frequency domain gaps in SR OCT images compared to the reference HR images, as they fail to resemble the high-frequency components, such as edges

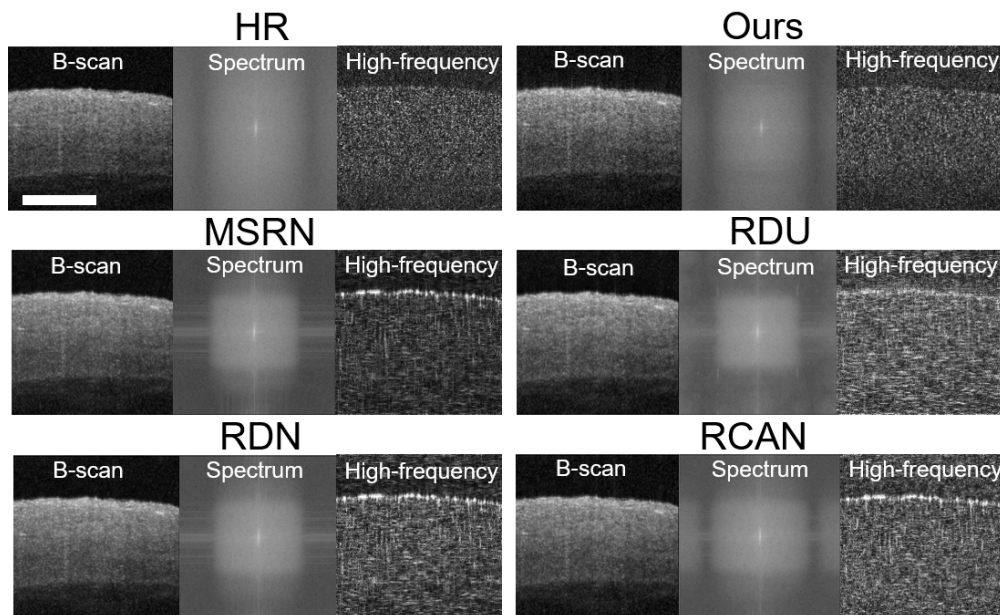


Fig. 1. Frequency domain gaps between the HR and the SR OCT images generated by four CNN implementations (MSRN, RDN, RDU, RCAN). The spectrum is generated by performing Fourier transform on the B-scan image. The high-frequency components of images are generated by performing an inverse Fourier transform on the high-frequency parts of the spectrum. Compared to the HR image, SR images generated by existing CNN algorithms are biased to a limited spectrum region towards low-frequency. Using ours CNN algorithm with frequency awareness (ours), the spectrum and high-frequency components of the SR OCT image are closer to that of the original image. The scale bar represents $500\mu\text{m}$.

and textures of the coronary artery sample. High-frequency components preserve finer details that are beneficial for medical imaging [32]. Recent works have used frequency information for reconstructing OCTA and natural images [33,34]. However, current research does not address the frequency domain gaps in the SR OCT images. Therefore, a DL framework with frequency awareness is needed to reduce spectral bias and generate high-quality SR OCT images.

To this end, we develop a deep learning framework that is capable of super-resolving LR OCT images with frequency awareness. In this manuscript, we propose a DL framework for OCT image SR task with frequency awareness, which is capable of restoring high-frequency information via model design and optimized loss function. We perform extensive experiments on an existing human coronary dataset and quantitatively demonstrate that the proposed frequency-aware DL framework super-resolves OCT images with superior quality and less frequency bias. We also validate the spectral bias of the existing DL algorithms used in generating SR OCT images. Furthermore, we perform qualitative analysis to confirm that our framework is capable of generating SR OCT images for corneal imaging and retinal imaging.

2. Methods

2.1. Overall framework

The design of our frequency-aware framework is shown in Fig. 2. Our framework consists of a generator (G) and a discriminator (D). Generator G translates a LR image to a SR image. Discriminator D classifies whether or not the generated image is realistic. Wavelet transformation is utilized to decompose feature maps F_i into different frequency components; frequency skip connection (FSC) is used to prevent the loss of high-frequency information; high-frequency alignment (HFA) is used to guide the G for generating frequency information [35].

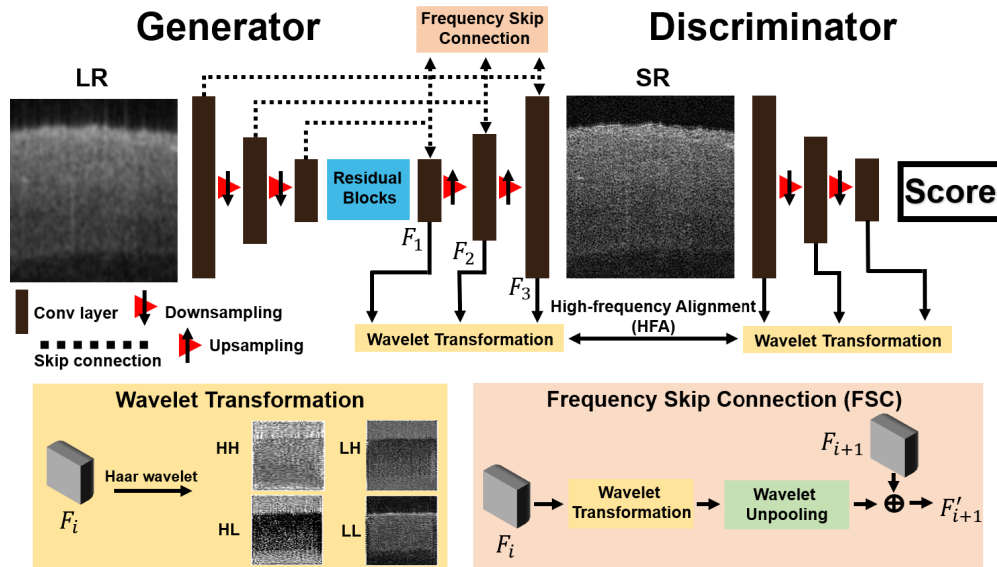


Fig. 2. The design of the proposed frequency-aware framework for OCT image super-resolution. The proposed model utilizes wavelet transformation, frequency skip connection, and high-frequency alignment to facilitate frequency information for super-resolving OCT images.

2.2. Model design

2.2.1. Wavelet transformation

We adopt Haar wavelet for decomposing feature maps of the i -th layer F_i into different components. Haar wavelet consists of two mirror operations: wavelet pooling and wavelet unpooling. The wavelet pooling converts images into the wavelet domain, and the wavelet unpooling inversely reconstructs frequency components into the spatial domain. During wavelet pooling, F_i is convolved with four distinct filters: LL^T , LH^T , HL^T , and HH^T , where L and H are low and high pass filters respectively ($L^T = \frac{1}{\sqrt{2}}[1, 1]$, $H^T = \frac{1}{\sqrt{2}}[-1, 1]$). The low pass filter (LL^T) provides general shapes and outlines in F_i ; the high pass filters (LH^T , HL^T , HH^T) provide more fine details such as segments, edges, and contrasts. An illustration of wavelet transformation is shown in Fig. 2.

2.2.2. Frequency skip connection

To prevent the loss of high-frequency information from F_i to F_{i+1} , FSC is used in generator G . The FSC in G is defined as:

$$F'_{i+1} = F_{i+1} + \text{Unpooling}(LL_G^i, LH_G^i, HL_G^i, HH_G^i) \quad (1)$$

After the frequency skip connection, feature map F'_{i+1} is calculated with the frequency information of F'_i is preserved through this process.

2.2.3. High-frequency alignment

High-frequency alignment (HFA) provides G with a self-supervised learning scheme using frequency information acquired in D . For F_i in G , we acquire LL_G^i , LH_G^i , HL_G^i , and HH_G^i . The combination of high-frequency components in G is defined by: $HF_G^i = LH_G^i + HL_G^i + HH_G^i$. Similarly, high-frequency components in D can be acquired by $HF_D^i = LH_D^i + HL_D^i + HH_D^i$. The HF_D^i can be used as the self-supervision constraint to train G .

2.3. Loss function

In the proposed frequency-aware framework, we incorporate a modified focal frequency loss (FFL) that quantifies the distance between HR and SR OCT images in the frequency domain [36]. The FFL is defined as:

$$FFL = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} w(u, v) |F_{SR}(u, v) - F_{HR}(u, v)|^2 \quad (2)$$

The F_{SR} and F_{HR} denote the frequency representation of SR and HR OCT images acquired by Discrete Fourier transform (DFT); the M and N represent the image size; the $w(u, v)$ is the spectrum weight matrix that is defined by:

$$w(u, v) = |F_{SR}(u, v) - F_{HR}(u, v)|^\alpha \quad (3)$$

where α is the scaling factor for flexibility (set to 1 in our experiments). In [36], the F_{SR} and F_{HR} are acquired by 2D DFT. However, the OCT images are acquired by 1D A-line scanning. Thus, we modify the FFL by acquiring F_{SR} and F_{HR} using 1D DFT. The original FFL is denoted as FFL_{2D} and the modified FFL is denoted as FFL_{1D} . The loss function L of the proposed

frequency-aware model is defined as:

$$\begin{aligned} L(G, D, L1, FFL_{1D}, FFL_{2D}, L_{align}) = \\ L_{adv}(G, D) + \alpha L1(SR, HR) + \beta FFL_{1D}(SR, HR) \\ + \beta FFL_{2D}(SR, HR) + \gamma L_{align}(G, D) \end{aligned} \quad (4)$$

The L_{adv} stands for the adversarial loss, which is defined by probabilities of the discriminator over generated images:

$$L_{adv} = -\log D(G(LR)) \quad (5)$$

The $L1$ stands for the mean absolute error, which is defined by the pixel difference between SR and HR images:

$$L_{L1} = |SR - HR| \quad (6)$$

The L_{align} stands for the distance of high-frequency information between the HR and SR OCT images:

$$L_{align} = \sum_{i=1}^3 |HF_D^i - HF_G^i| \quad (7)$$

We aim to solve the following minmax optimization problem:

$$G^* = \arg \min \max L(G, D, L1, FFL_{1D}, FFL_{2D}, L_{align}) \quad (8)$$

3. Results

3.1. Data collection

We perform a large-scale quantitative analysis on an existing coronary image dataset. The dataset was imaged using a commercial OCT system (Thorlabs Ganymede, Newton, NJ) [21]. The specimens were obtained in compliance with ethical guidelines and regulations set forth by Institutional Review Board (IRB), with de-identification to ensure the privacy of the subjects. A total of 2996 OCT images were obtained from 23 specimens, with a depth of 2.56 mm and a pixel size of $2 \mu\text{m} \times 2 \mu\text{m}$ within a B-scan. The width of the images varied from 2 mm to 4 mm depending on the size of the specimen.

In addition to the large-scale coronary dataset, we also confirmed the generalizability of applying the proposed method to two small datasets: one from *ex vivo* fish corneal and the other from *in vivo* rat retina. Two fish corneal OCT images were obtained from the same Thorlab OCT images following the same imaging protocol as coronary imaging. Fifty rat retinal images were obtained from a Heidelberg Spectralis SDOCT system. The system has an axial resolution of $7 \mu\text{m}$ and a lateral resolution of $14 \mu\text{m}$. The maximum field of view is 9 mm x 9mm. Animal imaging procedure was in accordance with protocols approved by the Institutional Animal Care and Use Committee at the Stevens Institute of Technology, and with the principles embodied in the statement on the use of animals in ophthalmic and vision research adopted by the Association for Research in Vision and Ophthalmology. The details of the experimental procedure are described in [37].

3.2. Experimental setup

The OCT LR images from coronary images and eye images were generated by cropping the spectrum data. The optical resolution of OCT systems will be decreased by reducing the bandwidth of the spectrum. We used $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{4}$ (denoted as X2, X3, and X4 respectively) of the raw spectrum data by central cropping. We used Hanning window to filter the raw spectrum data. Next, the filtered spectrum data were processed by FFT to get complex OCT data. Then, the norm of the complex OCT data was converted to dB scale. The background subtraction was

performed to remove noises in the OCT data. The LR OCT images were used as the inputs for the DL networks. The OCT images were randomly shuffled into five folds for cross-validation. The factors α , β , and γ in Eq. (4) are set to 10, 1, and 1 respectively.

3.3. Network implementation

We implemented our frequency-aware model using Pytorch. For downsampling layers, we used 2D convolutional layers with a stride of 2 followed by the Instance normalization layer and LeakyRelu activation layer with a negative slope of 0.2. For the upsampling layers, we used 2D transpose convolutional with a stride of 2 followed by the same Instance normalization and LeakyRelu activation. We used 16 residual blocks as the bottleneck, each containing 2 convolutional layers. For the implementation of previous DL algorithms (MSRN, RDN, RDU, RCAN), we were inspired by the designs in [21]. We implemented the existing DL algorithms in GAN architecture.

3.4. Training details

The image intensities were normalized to a range of [0,1]. The training protocol was performed five times for cross-fold validation. We randomly sampled 16 non-overlapping LR patches of size 64 x 64 pixels as input during training. The normalized images were augmented through random flipping to prevent overfitting. The optimization routine was carried out using the Adam algorithm, with an initial learning rate of 10^{-4} . A total of 200 epochs were executed to ensure convergence. The training process utilized one RTX A6000 GPU. Each training iteration on a single data fold consumed approximately 2 hours.

3.5. Evaluation metrics

We used peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [38] to measure the quality of SR images. The PSNR calculates pixel-wise differences between the HR image and the SR image, which is defined as:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (9)$$

The MSE represents the cumulative squared error between the HR and SR OCT images:

$$MSE = \frac{\sum_{m=1}^M \sum_{n=1}^N |HR(m, n) - SR(m, n)|^2}{M * N} \quad (10)$$

The SSIM focuses on structural similarity between the HR image and the SR image, which is defined as:

$$SSIM = \frac{(2\mu_{HR}\mu_{SR} + c_1)(2\sigma_{HR,SR} + c_2)}{(\mu_{HR}^2 + \mu_{SR}^2 + c_1)(\sigma_{HR}^2 + \sigma_{SR}^2 + c_2)} \quad (11)$$

where μ_{HR} is the pixel mean of the HR image; μ_{SR} is the pixel mean of the SR image; σ_{HR} is the variance of HR; σ_{SR} is the variance of SR; $\sigma_{HR,SR}$ is the covariance of HR and SR; $c_1 = (0.01 * 255)^2$ and $c_2 = (0.03 * 255)^2$ which are two variables to stabilize the division with weak denominator.

To evaluate the frequency difference, we define a frequency-level metric, namely Scaled Frequency Distance (SFD), which is defined as:

$$SFD = \sum_u^{M-1} \sum_v^{N-1} \left| \frac{|F_{SR}(u, v)| - |F_{HR}(u, v)|}{|F_{HR}(u, v)|} \right| \quad (12)$$

To evaluate the ability of our framework to preserve the edge details in the SR OCT images, we calculate the edge preservation index (EPI) of the region of interest (ROI). For EPI, 1 corresponds

to the ideal situation which means perfect edge preservation.

$$EPI = \frac{\sum_{m=1}^M \sum_{n=1}^N ((HR(m, n) - \text{mean}(HR)) * (SR(m, n) - \text{mean}(SR)))}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N (HR(m, n) - \text{mean}(HR))^2 * \sum_{m=1}^M \sum_{n=1}^N (SR(m, n) - \text{mean}(SR))^2}} \quad (13)$$

3.6. Analysis on spectral bias

We perform frequency analysis to evaluate the spectral bias of our frequency-aware model and other DL algorithms. We apply 2D DFT to the HR and SR OCT images, after which we average the logarithm of the intensities for each A-line and plot the intensity values over the pixels. The frequency analysis is carried out by averaging the spectrum of the SR OCT images. The results are reported in Fig. 3. As shown in Fig. 3 (a), our frequency-aware model generates SR images with averaged spectrums that are similar to the HR images. The summed intensities for pixels, as shown in Fig. 3 (b), confirm our frequency-aware model are less biased in spectrum distribution compared with other DL algorithms. On the other hand, existing DL algorithms generate SR OCT images with spectral bias in an unstable manner, as confirmed by Fig. 3.

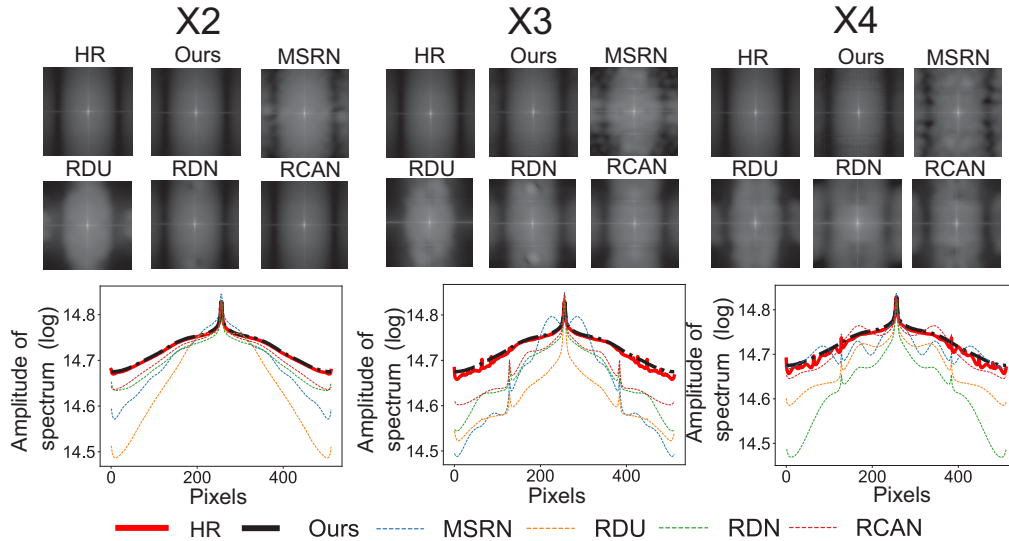


Fig. 3. Frequency analysis of the SR OCT images generated from LR OCT data acquired using factors of X2, X3, and X4. Compared to existing methods, our frequency-aware model is capable of super-resolving OCT images with less spectrum bias, which is confirmed by frequency analysis.

3.7. Quantitative analysis on super-resolution performance

We compare the quantitative performance of our frequency-aware model to other DL algorithms. As shown in Table 1, our frequency-aware model generates SR OCT images with better PSNR, SSIM, and SFD scores compared to other deep learning algorithms. Moreover, the standard deviation results show that our framework consistently outperforms existing methods. Together with Fig. 3, we confirm our frequency-aware model generates SR OCT images with better spatial and frequency properties compared to other DL algorithms.

In Fig. 4, a case of super-resolving an LR OCT image of a stent within the coronary artery is demonstrated. Coronary stent placement is an established treatment for CAD [39]. Imaging microstructures and tissues adjacent to stent struts are crucial in the clinic. It is critical to provide

Table 1. PSNR, SSIM, and SFD results of OCT images reconstructed by MSRN, RDU, RDN, RCAN, and our frequency-aware model. Red indicates the best performance. All results are averaged based on five-fold cross-validation. We report the mean value \pm standard deviation.

Methods	X2			X3			X4		
	PSNR \uparrow	SSIM \uparrow	SFD \downarrow	PSNR \uparrow	SSIM \uparrow	SFD \downarrow	PSNR \uparrow	SSIM \uparrow	SFD \downarrow
MSRN	30.2094 \pm 0.1256	0.8484 \pm 0.0030	0.3765 \pm 0.0022	24.2997 \pm 0.1898	0.4829 \pm 0.0100	0.8761 \pm 0.0578	23.6034 \pm 0.6719	0.4216 \pm 0.0250	0.7565 \pm 0.0247
RDU	29.1010 \pm 0.4972	0.7950 \pm 0.0370	0.6737 \pm 0.0473	23.8777 \pm 0.3115	0.4628 \pm 0.0080	0.8376 \pm 0.0124	23.5086 \pm 0.4143	0.4246 \pm 0.0230	0.9369 \pm 0.0626
RDN	30.3321 \pm 0.2031	0.8496 \pm 0.0050	0.3503 \pm 0.0090	23.7528 \pm 0.3553	0.4635 \pm 0.0160	0.6560 \pm 0.0682	23.7746 \pm 0.8651	0.4371 \pm 0.0280	0.6977 \pm 0.0456
RCAN	29.5280 \pm 0.7440	0.8261 \pm 0.0249	0.3918 \pm 0.0173	23.5904 \pm 0.5528	0.4596 \pm 0.0216	0.7638 \pm 0.0912	22.2740 \pm 0.4748	0.3462 \pm 0.0134	0.8121 \pm 0.0524
Ours	30.4713\pm0.2783	0.8519\pm0.0020	0.3273\pm0.0020	25.7500\pm0.0940	0.5904\pm0.0030	0.4669\pm0.0014	24.2867\pm0.0760	0.4424\pm0.0090	0.5287\pm0.0081

accurate morphological information on interactions between the stent and the vessel wall, for the purpose of evaluating the placement as well as the biocompatibility of the stent. The edges of the stent are considered to be high-frequency information in the OCT images, which are challenging to reconstruct for previous DL algorithms. As shown in Fig. 4, previous DL algorithms generate blurred edges of the stent. Moreover, existing DL algorithms lead to artifacts on the interaction between the stent and tissue, as shown in Fig. 4 (e), (f), (g). With our frequency-aware model, the edges of the stent are resolved with detailed information that is similar to that of the HR image. The EPI scores also show that the SR OCT image generated by our framework better preserves the edges.

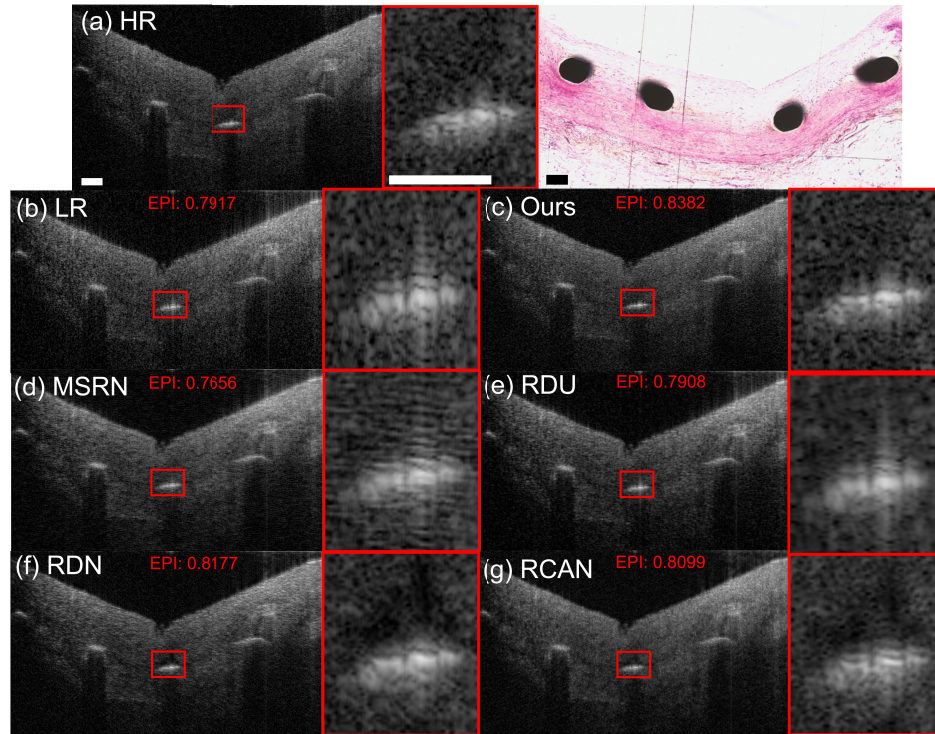


Fig. 4. Generating SR OCT images of stent structure from LR image acquired using a factor of X4. The corresponding histology image is attached. Our model resolves the boundary between the stent and tissue with better details due to its frequency-awareness design. ROIs are marked by red rectangles. The EPI score of the ROIs is calculated and displayed. The scale bar represents 100 μ m.

In Fig. 5, we demonstrate a case of super-resolving an LR OCT image of suspicious accumulation of macrophages. Macrophages play a critical role in both the development and rupture of atherosclerotic plaques [40], which are thus important for the diagnosis of CAD. OCT has been demonstrated to be a viable technique for visualizing the accumulation of macrophages in the human coronary artery. Macrophages appear as 'bright spots' in OCT images [41], which are high-frequency information due to their sharp contrast with neighboring tissues. As shown in red ROIs in Fig. 5, previous DL algorithms generate SR images with blurred macrophages, which will deteriorate the clinical diagnosis procedures. In contrast, our frequency-aware framework generates SR OCT images with clear macrophages. Thus, our frequency-aware model is capable of providing SR OCT images for human coronary samples with clinical meaningfulness. Also, we highlight edge regions in the orange ROIs. The EPI score suggests that our framework better preserves the edges in the SR OCT image.

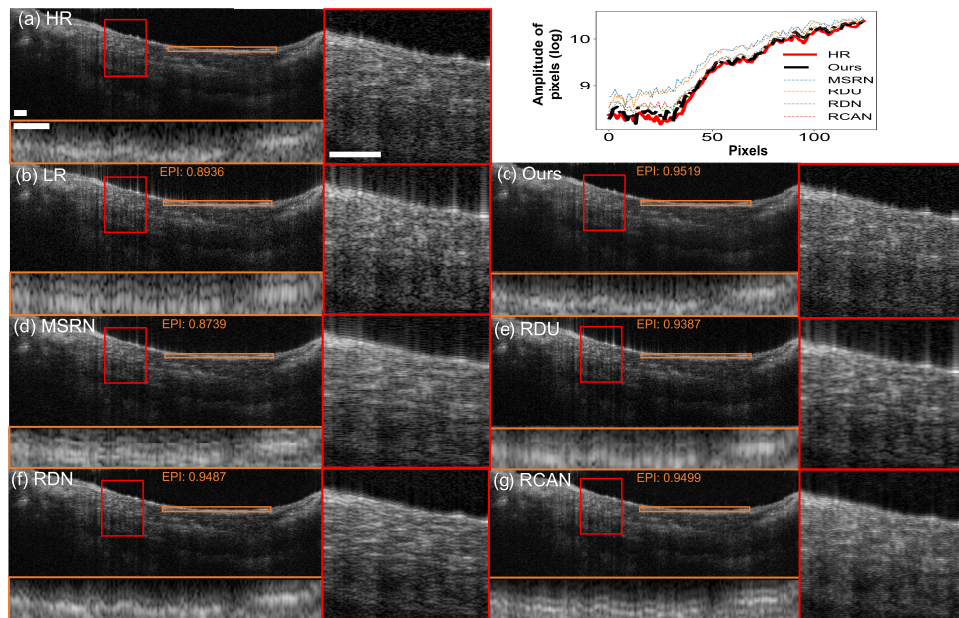


Fig. 5. Generating SR OCT image of suspicious macrophage regions from LR image acquired using a factor of X4. The amplitude of intensities of HR and SR images is attached. Our model resolves the accumulations of macrophages without blurring effects. ROIs containing the macrophage accumulations are marked by —red rectangles. ROIs of edge regions are marked by orange rectangles. The EPI score of the ROIs is calculated and displayed. The scale bar represents $100\mu\text{m}$.

3.8. Number of paired layers in model design

As shown in Fig. 2, we use three pairs of upsampling and downsampling layers in model design. As a comparative study, we design two variations of our framework, which use two and four pairs of upsampling and downsampling layers. We evaluate the performance of variations of our framework on the first data fold. The results are reported in Table 2. As shown in the table, our framework with three pairs of layers provides the best performance. With two pairs of layers, the framework may suffer from a lack of trainable parameters for performing the SR task, which leads to the sub-optimal performance. With four pairs of layers, the performance of the framework may still be limited because of the information flow bottleneck [42] in the UNet-like structures. On

Table 2. PSNR, SSIM, and SFD results of OCT images reconstructed by variations of our framework. Red indicates the best performance. The LR OCT images are generated using a factor of X2.

Metrics # of layers	PSNR↑	SSIM↑	SFD↓
2	30.1466	0.8406	0.3417
3	30.5339	0.8521	0.3153
4	30.0359	0.8352	0.3544

the other hand, the three pairs of upsampling/downsampling layers have been used in UNet-like structures relating to OCT images [43].

3.9. Weights of loss function

We further discuss the impact of three parameters, α , β , and γ in Eq. (4), on the performance. We adopt $\alpha = 10$, $\beta = 1$, and $\gamma = 1$ as the baseline. We also change the α to 100 and 1; β to 10 and 0.1; γ to 10 and 0.1. We evaluate the performance of the framework using loss function with different weights on the first data fold. The results are reported in Table 3. The results show that the baseline achieves the optimal performance, with two metrics of the best performance and one metric of the second-best performance.

Table 3. PSNR, SSIM, and SFD results of OCT images reconstructed by loss function with different weights. Red indicates the best performance. Blue indicates the second-best performance. The LR OCT images are generated using a factor of X2.

Metrics Weights	PSNR↑	SSIM↑	SFD↓
$\alpha = 10, \beta = 1, \gamma = 1$	30.5339	0.8521	0.3153
$\alpha = 100, \beta = 1, \gamma = 1$	30.5726	0.8468	0.3535
$\alpha = 1, \beta = 1, \gamma = 1$	21.5339	0.6636	0.3628
$\alpha = 10, \beta = 10, \gamma = 1$	30.0508	0.8319	0.3720
$\alpha = 10, \beta = 0.1, \gamma = 1$	30.0283	0.8321	0.3441
$\alpha = 10, \beta = 1, \gamma = 10$	29.5064	0.8166	0.3691
$\alpha = 10, \beta = 1, \gamma = 0.1$	29.2901	0.8278	0.3382

3.10. Application to super-resolve anterior segments of fish eye

Based on the setup in coronary imaging, we perform additional experiments on fish cornea using trained frequency-aware framework from previous section. We acquired the fish corneal OCT images using the same OCT system and imaging settings as the coronary dataset. We acquired three volumes of left and right fish eyes. Three representative OCT B-scans are used for the qualitative studies. The qualitative analysis of SR OCT images of fish corneal is shown in Fig. 6. In particular, the dash circle in the first panel shows the alignment of corneal stroma is better resolved after super-resolving. The dash circle in the second panel highlights the iris region that is underneath the corneal regions. The dash circle in the third panel resolves the bowman's layer in corneal region. Overall, our frequency-aware framework is capable of generating SR OCT fish corneal images with sharper and finer textures. Without retraining, our frequency-aware framework has the potential to be transferrable to OCT corneal images obtained from the same OCT system.

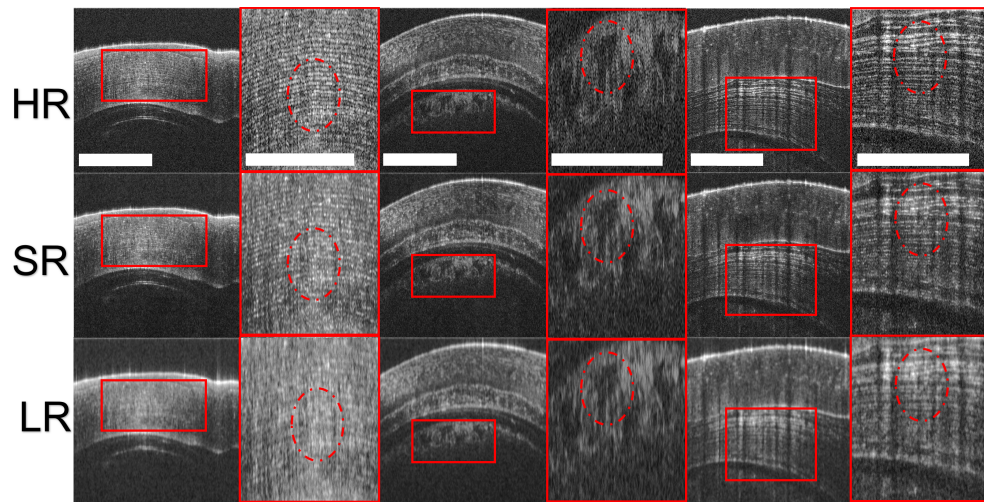


Fig. 6. Generating SR OCT images of anterior segments in fish eyes from LR images acquired using a factor of X3. ROIs are marked by red rectangles. The textures are highlighted by the dashed cycles. The scale bar represents $500\mu\text{m}$.

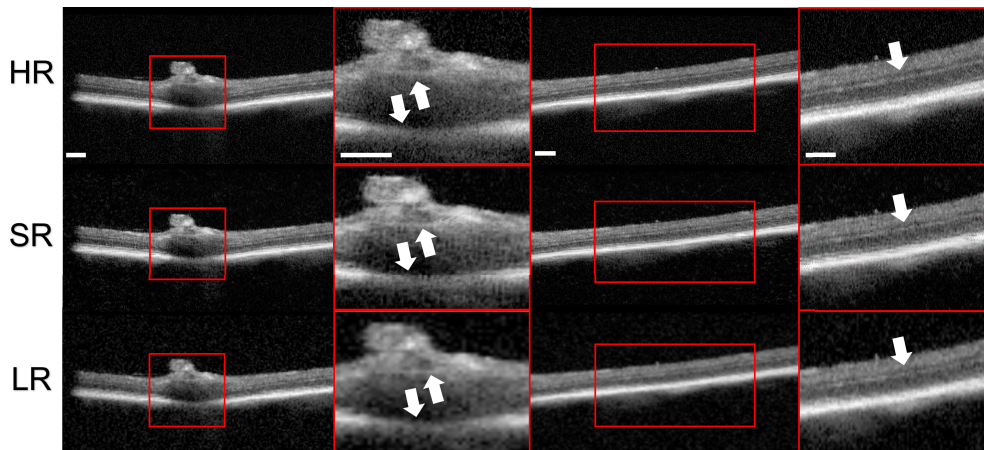


Fig. 7. Generating SR OCT images of posterior segments in rat eyes from LR images acquired using a factor of X3. ROIs are marked by red rectangles. The textures are highlighted by the white arrows. The scale bar represents $500\mu\text{m}$.

3.11. Application to super-resolve posterior segments of rat eye

We also conduct experiments to imaging posterior segments from animal model. A pigmented Long Evans rat from Charles River is used for OCT imaging. Retinal imaging requires a different optical design from benchtop OCT used in coronary imaging and corneal imaging. The OCT images were acquired using Heidelberg Spectralis system. We retrained the frequency-aware framework using 96% of the images and used the rest of the images for testing. The SR OCT images of rat eyes generated by our frequency-aware framework are shown in Fig. 7. In the first panel, the SR OCT image delineates the boundary around optical disc; and in the second panel, the SR OCT image better resolves the layer boundary within retinal regions (for example, inner nuclear layers). This experiment shows that our proposed frequency-aware framework, with

adequate retraining, has the potential to be generalizable to OCT retinal images obtained from different OCT systems.

4. Discussion

To the best of our knowledge, our framework is the first study in OCT community to propose a frequency-aware framework for super-resolution, which addresses the spectral bias in generated images. We design the proposed framework by modifying the convolution model architecture and loss functions to improve the frequency-aware capability. This is based on our investigation on the spectral bias towards the low-frequency components in the spectrum in existing studies. Our frequency-aware model generates SR OCT images with less spectral bias and better performance compared to existing framework. Our frequency-aware model is capable of generating SR OCT images with clinical meaningfulness, which is confirmed by qualitative analysis. Compared to recent research [33,34] which exploits frequency information for generating OCTA and natural images, our work exploits frequency information to reduce the spectral bias in SR OCT images, which is different from the OCTA and natural images. Also, our unified framework is differentiated from [29], which exploits frequency information of OCT images using multiple DL networks without addressing the spectral bias.

Another contribution lies in generalizability. Our preliminary study indicates great potential to be applied to multiple tissue types. We perform qualitative experiments on additional fish eye and rat eye dataset. Without retraining, our frequency-aware framework resolves anterior segments of fish eye, including corneal stroma, iris region, and downman's layer, acquired from the same OCT system. With adequate retraining, our frequency-aware framework is capable of resolving LR OCT images acquired from different systems. From the qualitative analysis of a rat eye dataset acquired from a different OCT system, we resolve the boundary around the optical disc and within retinal regions, with adequate retraining.

As an exploratory study on methodology development, this study, especially the eye imaging, is based on healthy samples. In the future, we plan to validate our super-resolution framework on pathological animal models to examine how much the improved resolution could facilitate the diagnosis and treatment in ophthalmology. Moreover, in addition to SDOCT, we plan to validate our super-resolution framework on swept-source OCT system in which the signal is also acquired in spectral domain. Also, we will extend and optimize the proposed framework for multiple tasks, such as OCT image denoising, classification, and segmentation.

5. Conclusion

In this paper, we investigate the spectral bias of existing DL algorithms when generating SR OCT images. To mitigate the spectral bias, we develop a frequency-aware model that combines cGAN with frequency loss to super-resolve LR OCT images. Compared to existing DL algorithms, our approach produces SR OCT images with less spectral bias, resulting in better preservation of textures. Additionally, our method generates SR coronary OCT images of superior quality, with higher PSNR and SSIM scores, as well as lower SFD scores. Our frequency-aware framework demonstrates the capability of generating SR OCT coronary images to provide better diagnosis and treatment. Moreover, our study also indicates the ability of the proposed framework to be generalized to corneal imaging and retinal imaging.

Funding. National Eye Institute (R01EY029298, R01EY032222); New Jersey Health Foundation; National Science Foundation (2222739, 2239810).

Acknowledgement. The authors would like to thank Chaimae Gouya, Mohammed Attia, and Aaron Shamouil for their assistance in OCT data acquisition.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. Y. Ling, Z. Dong, X. Li, Y. Gan, and Y. Su, "Deep learning empowered highly compressive ss-OCT via learnable spectral-spatial sub-sampling," *Opt. Lett.* **48**(7), 1910–1913 (2023).
2. A. Al-Mujaini, U. K. Wali, and S. Azeem, "Optical coherence tomography: clinical applications in medical practice," *Oman Med. J.* **28**(2), 86–91 (2013).
3. X. Li, H. Liu, X. Song, B. C. Brott, S. H. Litovsky, and Y. Gan, "Structural constrained virtual histology staining for human coronary imaging using deep learning," (2022).
4. H. Liu, X. Li, A. L. Bamba, X. Song, B. C. Brott, S. H. Litovsky, and Y. Gan, "Toward reliable calcification detection: calibration of uncertainty in object detection from coronary optical coherence tomography images," *J. Biomed. Opt.* **28**(03), 036008 (2023).
5. S.-Y. Lee and M.-K. Hong, "Stent evaluation with optical coherence tomography," *Yonsei Med. J.* **54**(5), 1075–1083 (2013).
6. Y. Yang, E. Pavlatos, W. Chamberlain, D. Huang, and Y. Li, "Keratoconus detection using OCT corneal and epithelial thickness map parameters and patterns," *J. Cataract Refract. Surg.* **47**(6), 759–766 (2021).
7. A. Kamalipour and S. Moghimi, "Macular optical coherence tomography imaging in glaucoma," *J. Ophthalmic & Vis. Res.* **16**(3), 478–489 (2021).
8. D. Scuteri, A. Vero, M. Zito, M. D. Naturale, G. Bagetta, C. Nucci, P. Tonin, and M. T. Corasaniti, "Diabetic retinopathy and age-related macular degeneration: a survey of pharmacoutilization and cost in calabria, italy," *Neural Regener. Res.* **14**(8), 1445–1448 (2019).
9. E. Midena, T. Torresin, E. Velotta, E. Pilotto, R. Parrozzani, and L. Frizziero, "OCT hyperreflective retinal foci in diabetic retinopathy: A semi-automatic detection comparative study," *Front. Immunol.* **12**, 1 (2021).
10. B. L. Sikorski, G. Malukiewicz, J. Stafiej, H. Lesiewska-Junk, and D. Raczynska, "The diagnostic function of OCT in diabetic maculopathy," *Mediators Inflammation* **2013**, 1–12 (2013).
11. Y. Liu, Y. Liang, G. Mu, and X. Zhu, "Deconvolution methods for image deblurring in optical coherence tomography," *J. Opt. Soc. Am. A* **26**(1), 72–77 (2009).
12. S. A. Hojjatoleslami, M. R. N. Avanaki, and A. G. Podoleanu, "Image quality improvement in optical coherence tomography using lucy-richardson deconvolution algorithm," *Appl. Opt.* **52**(23), 5663–5670 (2013).
13. Y. Chen, J. Fingler, and S. E. Fraser, "Multi-shaping technique reduces sidelobe magnitude in optical coherence tomography," *Biomed. Opt. Express* **8**(11), 5267–5281 (2017).
14. X. Liu, S. Chen, D. Cui, X. Yu, and L. Liu, "Spectral estimation optical coherence tomography for axial super-resolution," *Opt. Express* **23**(20), 26521–26532 (2015).
15. Y. Xu, B. M. Williams, B. Al-Bander, Z. Yan, Y. Shen, and Y. Zheng, "Improving the resolution of retinal OCT with deep learning," in *Annual Conference on Medical Image Understanding and Analysis*, (2018).
16. K. Liang, X. Liu, S. Chen, J. Xie, W. Q. Lee, L. Liu, and H. K. Lee, "Resolution enhancement and realistic speckle recovery with generative adversarial modeling of micro-optical coherence tomography," *Biomed. Opt. Express* **11**(12), 7236–7252 (2020).
17. Q. Hao, K. Zhou, J. Yang, Y. Hu, Z. Chai, Y. Ma, G. Liu, Y. Zhao, S. Gao, and J. Liu, "High signal-to-noise ratio reconstruction of low bit-depth optical coherence tomography using deep learning," *J. Biomed. Opt.* **25**(12), 123702 (2020).
18. A. Lichtenegger, M. Salas, A. Sing, M. Duell, R. Licandro, J. Gesperger, B. Baumann, W. Drexler, and R. A. Leitgeb, "Reconstruction of visible light optical coherence tomography images retrieved from discontinuous spectral data using a conditional generative adversarial network," *Biomed. Opt. Express* **12**(11), 6780–6795 (2021).
19. B. Qiu, Y. You, Z. Huang, X. Meng, Z. Jiang, C. Zhou, G. Liu, K. Yang, Q. Ren, and Y. Lu, "N2nsr-OCT: Simultaneous denoising and super-resolution in optical coherence tomography images using semisupervised deep learning," *J. Biophotonics* **14**(1), e202000282 (2021).
20. Y. Zhang, T. Liu, M. Singh, E. Çetintaş, Y. Luo, Y. Rivenson, K. V. Larin, and A. Ozcan, "Neural network-based image reconstruction in swept-source optical coherence tomography using undersampled spectral data," *Light: Sci. Appl.* **10**(1), 155 (2021).
21. X. Li, S. Cao, H. Liu, X. Yao, B. C. Brott, S. H. Litovsky, X. Song, Y. Ling, and Y. Gan, "Multi-scale reconstruction of undersampled spectral-spatial OCT data for coronary imaging using deep learning," *IEEE Trans. Biomed. Eng.* **69**(12), 3667–3677 (2022).
22. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds. (Springer International Publishing, Cham, 2015), pp. 234–241.
23. Y. Huang, Z. Lu, Z. Shao, M. Ran, J. Zhou, L. Fang, and Y. Zhang, "Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network," *Opt. Express* **27**(9), 12289–12307 (2019).
24. Y. Huang, W. Xia, Z. Lu, Y. Liu, H. Chen, J. Zhou, L. Fang, and Y. Zhang, "Noise-powered disentangled representation for unsupervised speckle reduction of optical coherence tomography images," *IEEE Trans. Med. Imaging* **40**(10), 2600–2614 (2021).

25. M. Geng, X. Meng, L. Zhu, Z. Jiang, M. Gao, Z. Huang, B. Qiu, Y. Hu, Y. Zhang, Q. Ren, and Y. Lu, "Triplet cross-fusion learning for unpaired image denoising in optical coherence tomography," *IEEE Trans. Med. Imaging* **41**(11), 3357–3372 (2022).
26. Z. Yuan, D. Yang, H. Pan, and Y. Liang, "Axial super-resolution study for optical coherence tomography images via deep learning," *IEEE Access* **8**, 204941–204950 (2020).
27. H. Pan, D. Yang, Z. Yuan, and Y. Liang, "More realistic low-resolution OCT image generation approach for training deep neural networks," *OSA Continuum* **3**(11), 3197–3205 (2020).
28. V. Das, S. Dandapat, and P. K. Bora, "Unsupervised super-resolution of OCT images using generative adversarial network for improved age-related macular degeneration diagnosis," *IEEE Sens. J.* **20**(15), 8746–8756 (2020).
29. W. Lee, H. S. Nam, J. Y. Seok, W.-Y. Oh, J. W. Kim, and H. Yoo, "Deep learning-based image enhancement in optical coherence tomography by exploiting interference fringe," *Commun. Biol.* **6**(1), 464 (2023).
30. N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, "On the spectral bias of neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, vol. 97 of *Proceedings of Machine Learning Research* K. Chaudhuri and R. Salakhutdinov, eds. (PMLR, 2019), pp. 5301–5310.
31. M. Tancik, P. P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, (Curran Associates Inc., 2020), NIPS'20.
32. D. McIntosh, T. P. Marques, and A. B. Albu, "Preservation of high frequency content for deep learning-based medical image classification," in *2021 18th Conference on Robots and Vision (CRV)*, (2021), pp. 41–.
33. W. Zhang, D. Yang, C. Y. Cheung, and H. Chen, "Frequency-aware inverse-consistent deep learning for OCT-angiogram super-resolution," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, eds. (Springer Nature Switzerland, Cham, 2022), pp. 645–655.
34. J. Li, T. Dai, M. Zhu, B. Chen, Z. Wang, and S.-T. Xia, "FSR: A general frequency-oriented framework to accelerate image super-resolution networks," *Proc. AAAI Conf. on Artif. Intell.* **37**(1), 1343–1350 (2023).
35. M. Yang, Z. Wang, Z. Chi, and Y. Zhang, "FreGAN: Exploiting frequency components for training gans under limited data," in *Advances in Neural Information Processing Systems*, vol. 35 S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, eds. (Curran Associates, Inc, 2022), pp. 33387–33399.
36. L. Jiang, B. Dai, W. Wu, and C. C. Loy, "Focal frequency loss for image reconstruction and synthesis," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, (2021), pp. 13899–13909.
37. W. Liu, A. P. Tawakol, K. M. Rudeen, W. F. Mieler, and J. J. Kang-Mieler, "Treatment efficacy and biocompatibility of a biodegradable aflibercept-loaded microsphere-hydrogel drug delivery system," *Trans. Vis. Sci. Tech.* **9**(11), 13 (2020).
38. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
39. S. Elezi, A. Kastrati, F.-J. Neumann, M. Hadamitzky, J. Dirschinger, and A. Schömig, "Vessel size and long-term outcome after coronary stent placement," *Circulation* **98**(18), 1875–1880 (1998).
40. T. J. Barrett, "Macrophages in atherosclerosis regression," *Arterioscler., Thromb., Vasc. Biol.* **40**(1), 20–33 (2020).
41. G. J. Tearney, "OCT imaging of macrophages: a bright spot in the study of inflammation in human atherosclerosis," *JACC. Cardiovasc. Imaging* **8**(1), 73–75 (2015).
42. S. Lee and I. V. Bajić, "Information flow through u-nets," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (2021), pp. 812–816.
43. J. Kugelman, J. Allman, S. A. Read, S. J. Vincent, J. Tong, M. Kalloniatis, F. K. Chen, M. J. Collins, and D. Alonso-Caneiro, "A comparison of deep learning u-net architectures for posterior segment OCT retinal layer segmentation," *Sci. Rep.* **12**(1), 14888 (2022).